JOURNAL OF ECONOMIC THEORY 54, 124–147 (1991)

On the Competitive Pressure Created by the Diffusion of Innovations*

DILIP MOOKHERJEE AND DEBRAJ RAY

Indian Statistical Institute, New Delhi 110016, India

Received August 30, 1988; revised September 17, 1990

We consider the decision of a dominant firm to adopt a sequence of potential cost-reducing innovations, where the latest technology adopted diffuses to a competitive fringe at an exogenous rate. With price competition on the product market, the leader optimally spaces apart the adoption dates of successive innovations, so the industry is characterized by Schumpetarian cycles of alternating innovation and diffusion. An increase in the rate of diffusion has ambiguous effects on innovative activity, and, up to a point, hastens the pace of innovation. These results may, however, be reversed in the case of quantity competition. *Journal of Economic Literature* Classification Numbers: 022, 612, 621.

1. INTRODUCTION

Recent years have witnessed the development of a large literature analyzing the incentives for firms to carry out R & D, and adopt cost-saving innovations (see, e.g., Arrow [2], Leibenstein [10], Kamien and Schwartz [7–9], Loury [11], Dasgupta and Stiglitz [3, 4], Reinganum [14], and Spence [19]). One of the central insights of this literature deals with the tension between the requirement of static allocative efficiency subsequent to an innovation, and that of dynamic incentives for innovation *ex ante*. Due to inherent inappropriabilities and spillovers, an innovator cannot obtain the full social surplus associated with an innovation. In fact, if the innovation already exists, it is socially desirable that it diffuses as rapidly as possible. But this would leave potential innovators with no *ex ante* incentives to expend resources on innovative activity.

This insight can be traced back to Schumpeter [16, 17], who emphasized that in the long run it may well be more important to promote *ex ante* innovation incentives, at the expense of restricting the diffusion of

* We are grateful to an associate editor for outlining an argument that greatly simplified the proofs of the results, and for detailed expositional advice that has improved the quality of the paper significantly. We also thank two anonymous referees for useful comments. This paper is dedicated to Angela Bhaya Soares and her abiding interest in Schumpeter.

innovation *ex post.*[1] The empirical work of Abramovitz [1], Denison [5], and Solow [18] has further heightened the importance of technical progress as a key determinant of economic growth. Consequently, economists have been reluctant to advise weakening of patent protection or the pursuit of aggressive procompetitive policies in industries with significant potential for technical progress.

An alternative viewpoint on the impact of competition on innovation incentives, associated in particular with the work of Harvey Leibenstein, is that competition acts as a "stick" that promotes innovative activity.[2] From this perspective, it could be argued that faster diffusion of innovations enables followers to catch up faster with the innovator, thereby enhancing the latter's incentive to introduce the next set of innovations. Such a viewpoint necessitates a dynamic formulation of the innovation process, i.e., as one comprising a *sequence* of innovations. Alternatively, one may consider a single innovation with a variable date of adoption. In this paper, we seek to investigate the validity of this intuition in alternative settings.

Most theoretical literature on innovation incentives, however, employs a static formulation of the innovation process; it usually analyzes the incentive for a firm to adopt a single innovation instantaneously or never at all.[3] No one would dispute that in reality, technical change is associated with a sequence of multiple innovations. Indeed, Schumpeter's theory of capitalist development placed central emphasis on the cyclical nature of the innovation process, where successive cycles are associated with different innovations. In such dynamic settings, the intensity of innovative activity is necessarily multidimensional. For instance, the total *number* of innovations adopted within a given time period provides one aspect of the extent of innovative activity; another is the *timing* of adoption decisions. The case where a single innovation is never adopted may be considered almost equivalent to the case where one is adopted in the very distant future. Alternatively, it may be socially preferable to have a few innovations adopted instantly, rather than a larger number of innovations adopted at some distant future date.

The purpose of this paper is twofold. First, we develop a model of a dynamic innovation process with some "Schumpetarian" elements. Our model permits alternating cycles of innovation and diffusion. Second, we

[1] To be sure, Schumpeter [17] did emphasize the notion of (dynamic) competition, and its importance in the innovation process: this is a crucial ingredient of his concept of "creative destruction." However, Chapter VIII of "Capitalism, Socialism, and Democracy" contains a clear statement of the virtues of monopoly in protecting the rents of innovators, thereby generating *ex ante* innovation incentives.

[2] See Leibenstein [10], especially Section III.

[3] Exceptions are Futia [6] and Reinganum [14]. Also, some interesting simulation experiments can be found in Nelson and Winter [13].

investigate the effects of an exogenous increase in the diffusion rate, on both the number of innovations adopted, as well as their timing. We use the simplest possible formulation that permits an assessment of the relative validity of the two contrasting viewpoints discussed above.

Our model has a single dominant firm with access to a sequence of potential innovations. The adoption of any innovation necessitates a fixed cost, to be incurred at the date of innovation.[4] Each innovation enables an immediate reduction in the unit cost of production of a homogeneous good or service. This can be interpreted to incorporate both process and product innovations.[5]

Follower firms are assumed unable to innovate on their own. Instead, they gradually imitate the innovations introduced by the leader. This allows their unit costs to drift down over time, approaching the leader's current costs at an exogenous rate λ. At each date, all firms simultaneously choose prices or quantities of their respective products. We adopt the simplifying assumption that payoffs on the product market are given by a single-period equilibrium of the relevant price (or quantity) setting game—there is no tacit collusion. Since the good is homogeneous, equilibrium in the price game is characterized by absence of actual production by the follower firm: the leader prevents entry by pricing below the follower's cost. In contrast, the quantity game involves positive market shares for both firms.

In the context of this model, we pose the following questions:

(i) Under what conditions is it optimal for the dominant firm to space apart the adoption dates of successive innovations, rather than carry them out simultaneously, or in close proximity to each other? This issue has a bearing on Schumpeter's contention that the innovation process is inherently cyclical.

(ii) What is the effect of an increase in the diffusion rate λ on the *ex ante* incentives of the leader to adopt innovations? In this context, no less important than the effects on the number and volume of these innovations are the implications for their *timing*. The answer to this question should have implications for government policies that affect the severity of patent protection (which influences the diffusion rate λ).

[4] We concentrate on the adoption of methods based on existing knowledge, rather than R & D aimed at producing new knowledge, which involves development costs over an entire development period prior to the date of invention. This aspect, as well as our interest in a dynamic sequence of innovations, and most important, our emphasis on the nature of product market competition, differentiate our model from the interesting work of Kamien and Schwartz [8, 9].

[5] See Spence [19] for a description of how product innovations could be viewed as a reduction in the costs of providing services to customers. See also Schumpeter [17, p. 92].

Our results indicate that the answers to the above questions depend crucially on the nature of product market competition. Consider question (i). We show that in the quantity-setting model, the adoption dates of available innovations will be bunched together towards the beginning. The leader thus moves away from the competitive fringe and continues to do so in the early period of the industry. A period of early dominance results, which is then followed by a phase of inter-firm cost equalization as the innovations diffuse to the competitive fringe. However, the price competition model yields very different results. The existing stock of innovations are adopted in a phased manner, and the industry displays alternate periods of innovation and diffusion. So far as the adoption of existing knowledge is concentrated, the Schumpetarian cyclical process seems to require competition in prices rather than quantities.[6]

We turn now to question (ii). As mentioned above, the existing theoretical literature has largely focused on a static notion of innovation, where only one potential innovation is available, to be adopted "now" or not at all. In such a context, increasing the diffusion rate λ will unambiguously reduce innovation incentives. But with more than one potential innovation (or a variable date of adoption of a single innovation), matters are more complex. Consider the price-setting model: there, an increase in the diffusion rate cannot increase the total number of innovations adopted. But the effect on the *timing* of these innovations is ambiguous. Over some range of values of the diffusion rate, an increase in this rate will advance the adoption date of the next innovation, without lowering the total number of innovations adopted. Over other ranges, it may delay adoption and/or reduce the total number adopted. The former result is explained by the Leibensteinian intuition that an increase in the diffusion rate enables followers to catch up faster, increasing the competitive pressure on the leader and motivating quicker adoption of further innovations.

Price competition is essential to this result. It is this form of interaction in the product market that induces innovation through competitive pressure. Indeed, this is precisely why question (i) is answered the way it is: the continuing adoption of innovations is motivated by a drop in follower costs, for which a minimum interval of time must elapse. Despite the fact that follower firms never actually produce, their *potential* ability to produce at progressively lower cost levels exerts competitive pressure on the leader.

[6] Of course, if a new stock of innovations becomes available, the quantity-setting model might also display cyclical phases of innovation. However, these phases would be correlated with the arrival of new knowledge. This is different from the periodic innovations predicted by the price-setting model.

On the other hand, we will see that in the quantity competition model innovations are encouraged by the relative absence of competitive pressure. This is why the leader must adopt an innovation early if he adopts it at all. Since all innovations are bunched together at the very beginning, an increase in the diffusion rate reduces the number of innovations adopted, without affecting their timing.

Our results therefore capture the contrasting views of Schumpeter and Leibenstein regarding the effect of competition on innovation incentives. In the quantity competition model, the forces envisaged by Leibenstein have minimal impact, and the classical Schumpetarian thesis regarding the negative role of diffusion on innovation is vindicated. However, as described above, matters are quite different in the price competition model: some intermediate value of the diffusion rate may well be socially optimal for overall incentives to innovate. And to this extent, the tension between static efficiency and dynamic incentives is less than what it is commonly supposed to be. A certain degree of imitation and spillover not only reduces the static allocative inefficiencies associated with the monopoly power of the innovator, but may also increase the incentives for rapid innovation in the future.

This paper is organized as follows. Section 2 introduces some examples to illustrate the basic intuition underlying our analysis. Section 3 then describes the general framework and Section 4 considers the price competition model in detail. The contrasting model of quantity competition is treated in Section 5. Some of the longer proofs are relegated to the Appendix.

2. Examples

In this section, we discuss in an intuitive fashion some of the issues intrinsic to a dynamic innovation model. Consider the simplest possible scenario where a pair of duopolists compete on a market for a homogeneous product. The firms have constant marginal cost; firm 1's cost is c and firm 2's is r, where $0 < c < r$. Call firm 1 the leader, and firm 2 the follower. Denote the leader's reduced form profit function on the product market by $L(r, c)$, where L is increasing in r and decreasing in c.

Suppose initially that the leader considers a single innovation which reduces his unit cost from c to c'. The follower's cost also drops by a fraction of the cost reduction of the leader: r drops to $r' = r - \lambda(c - c')$, where λ represents the extent of diffusion. This is akin to the standard formulation of diffusion in static models of innovation, as in Spence [19]. The leader's gross benefit from the innovation is equal to $[L(r', c') - L(r, c)]$. Clearly, an increase in the diffusion rate λ reduces the follower's post-innovation cost r', and thereby diminishes the leader's incentive to innovate. This is the basic Schumpetarian effect.

Now suppose that the leader has available a sequence of two potential innovations. The first innovation is as described above: conditional on adopting it, the leader may consider adopting the second one which will reduce his cost from c' to c'', and the follower's cost from r' to $r''=r'-\lambda(c'-c'')$. The leader's benefit from the second innovation is $[L(r'', c'')-L(r', c')]$. Since an increase in the diffusion rate λ lowers both r'' and r', the effect on innovation incentives is not so clear any more. It is conceivable that the effect on r' outweighs that on r'', i.e., the fact that the follower catches up more since the first innovation (the "Leibensteinian" stick) has a stronger effect than the reduction of post-innovation rents $L(r'', c'')$ (the "Schumpetarian" carrot).

The presence of a potential sequence of innovations is a major feature of our analysis below. A second important feature is the explicit role of product market competition, to which we now turn.

To form intuition about the importance of the nature of product market competition, consider the artificial "static" scenario where there is no diffusion and the follower's cost is set at r, both before and after the innovation. The leader's benefit from innovation is $E(r)\equiv L(r, c')-L(r, c)$.

Now ask the question: Does a decrease in r increase the innovation incentive $E(r)$? When both pre-innovation and post-innovation cost levels of the follower are decreased, both $L(r, c')$ and $L(r, c)$ are reduced. In this sense both the "carrot" and "stick" co-exist in this thought experiment, though in a simplified form somewhat different from the "real" situation.

The behavior of $E(r)$ depends on the nature of product market competition. The answer obviously depends on the sign of the cross partial derivative $L_{rc}(r, c)$. To illustrate most simply, assume that the product market is characterized by a linear demand curve

$$P=\lambda-\beta q, \qquad \lambda, \beta>0, \tag{1}$$

where P denotes price and q the sum of outputs of the two firms.

In the Cournot case, assuming positive production for each firm, routine computation shows that the leader's profit under the cost configuration (c, r) is given by $(\lambda-2c+r)^2/9\beta$. Consequently,

$$E(r)=\frac{(\lambda-2c'+r)^2}{9\beta}-\frac{(\lambda-2c+r)^2}{9\beta} \tag{2}$$

so that

$$E'(r)=\frac{4(c-c')}{9\beta}>0. \tag{3}$$

Therefore, *under quantity competition, the sign of L_{rc} is positive, and a lower follower cost reduces the incentive to innovate*, ceteris paribus.

Turn now to the case of price competition. For the follower's cost to matter at all, it must be the case that the leader limit prices the follower.[7] Therefore, the leader's profit under the cost configuration (c, r) (with $c < r$) is $((\lambda - r)/\beta)(r - c)$, so that

$$E(r) = \frac{\lambda - r}{\beta}(c - c'). \tag{4}$$

Therefore

$$E'(r) = -\frac{c - c'}{\beta} < 0. \tag{5}$$

Under price competition, the sign of L_{rc} is negative, and a lower follower cost increases the incentive to innovate.

These observations are easily understood using the Envelope Theorem, which states that the effect of cost reduction on the leader's profit is proportional to the quantity of output produced by the leader. In the presence of potential price competition from the follower, the leader is forced to "limit-price" at slightly below the follower's cost. A reduction in the latter resulting from more diffusion then causes the leader to set a lower price. Since the demand curve is downward sloping this increases the leader's output, and therefore also increases the benefit from a cost-reducing innovation. In contrast, a reduction in the follower's cost in the case of quantity competition causes the leader's output to decline.

An anonymous referee has pointed out that the preceding result for the price competition model may not hold if the goods produced by the two firms are differentiated rather than homogeneous. For instance, if the demand function facing firm i is $q_i = \alpha - \beta p_i + \gamma p_j$, with $\beta > \gamma$ and $j \neq i$, then it can be verified that for interior equilibria $E'(r) > 0$. In this context, the follower produces a positive quantity in equilibrium, unlike the homogeneous good context. This may lead one to conjecture that the crucial characteristic distinguishing our two models of product market competition is actual versus potential competition, rather than the strategic variables chosen (i.e., prices versus quantities). We do not, however, consider the case of differentiated products in this paper, and the validity of this conjecture remains an important open question.

The main results of our paper are based on the fact that the informal line of reasoning outlined in preceding paragraphs continues to hold in an

[7] For further discussion of this point, see Footnote 9 and the accompanying discussion below.

explicitly dynamic model, where the leader has access to a sequence of cost-reducing innovations, and where the follower's cost drifts towards that of the leader's. It turns out that there are a number of additional complexities involved in establishing the desired results, stemming from the dynamic nature of the analysis. Here is an example of one such difficulty.

Consider the price-setting formulation, and suppose that the follower's cost is diffusing towards that of the leader, at an exponential rate λ, starting from some initial value $r > c$. The process is

$$\dot{r}'_t = -\lambda(r'_t - c'), \qquad r'_0 = r \tag{6}$$

if the leader innovates, and

$$\dot{r}_t = -\lambda(r_t - c), \qquad r_0 = r \tag{7}$$

if it does not adopt the innovation. The gain from the innovation at date 0, expressed in present-value terms under a discount rate $\rho > 0$ is then

$$E(r) = \int_0^\infty e^{-\rho t} \left[\frac{\lambda - r'_t}{\beta} (r'_t - c') - \frac{(\lambda - r_t)}{\beta} (r_t - c) \right] dt. \tag{8}$$

Taking derivatives and using (6) and (7), we have

$$E'(r) = -\frac{1}{\beta} \int_0^\infty e^{-(\rho + \lambda) t} [(c - c') - 2(r_t - r'_t)] \, dt. \tag{9}$$

The "static" analysis applied to (9), only applies at date 0, when $r_t = r'_t$. The point, however, is that *over time*, the cost of the follower is *not* independent of whether or not the leader had adopted the innovation. Indeed, the sign of the expression within the integral of (9) depends on the date t; it can be shown to be initially negative, and eventually strictly positive (as $t \to \infty$). Nevertheless the *entire* integral can be shown to be negative. These results are shown in succeeding sections to extend to the case of an arbitrary (finite) number of innovations, with adoption dates chosen endogenously, as well as to all downward-sloping demand curves.

Our purpose in this section has been twofold. First, we discussed an example to provide the basic intuition driving our analysis, differentiating, in particular, between the quantity and price-setting models. Second, we argued that there are further considerations that are inherently dynamic, so that the static intuition does not take us "all the way." The proofs of our main results are consequently somewhat complicated. Nevertheless, a dynamic model is indispensable for the kind of questions we are interested in.

3. The Model

There is a single leader and a "competitive fringe" of followers, which we shall represent by a single firm. There is available (to the leader) a finite stock of cost-reducing innovations, numbered 1 to n. The innovations must be adopted in the given order, though two or more adjacent innovations can be adopted simultaneously. It will help to think of this stock as an existing body of scientific knowledge capable of improving the product or the production process, which can be adopted in *stages*. Accordingly, we shall often refer to each innovation in the sequence as a stage. The leader is said to be *in stage k* if he has adopted the first k innovations. So the first stage is 0, and the last is n.

The ith innovation costs an amount X^i to adopt, and this cost is incurred at the time of innovation. An innovation has the effect of instantly reducing the unit production cost of the innovator. We therefore associate with the ith stage a unit cost c^i, where $c^0 > c^1 > \cdots > c^n > 0$. The initial cost of the leader is c^0, while that of the follower is r^0, and we assume that $r^0 > c^0$.

We adopt the Schumpetarian assumption that the follower lacks the ability to innovate. However, the innovations adopted by the leader diffuse to the follower, creating a downward drift in the follower's unit cost. Specifically, assume that if at time t, the follower's cost is r_t, and the leader's cost is c_t (with $r_t > c_t$), then

$$\dot{r}_t = -\lambda(r_t - c_t), \qquad \text{with} \quad r_0 = r^0 \tag{10}$$

for some $\lambda > 0$. This parameter λ is the *diffusion rate*.

At any given instant of time, the two firms compete in the product market. We assume that there is no tacit collusion, so that we can write the one-period profit of the leader solely as a function of the *current* vector of costs: $L(r, c)$.[8]

One cannot hope for a general theory that applies irrespective of the nature of product market behavior, as subsumed in the form of the payoff function L. Indeed, this is a major point that we shall emphasize. In particular, we shall be contrasting the cases of price competition and quantity competition. We now proceed to define these two models.

[8] It can be checked that the analysis generalizes to the case where there are a large number of followers, all of whom have the same cost r_t (at time t). In fact, in the price-setting case, our analysis is consistent with free entry (i.e., an infinite number of potential firms, all of whom have cost r_t). Consequently, the absence of tacit collusion is a natural assumption in these contexts. In the quantity-setting model, however, the generalization to more than one follower requires more work.

Let $D(p)$ be the aggregate demand curve for the product. Throughout, we shall assume that D is differentiable and strictly decreasing, with $D'(p) < 0$. Further, we assume that the profit function $\pi(p, c) = (p-c)\,D(p)$ attains a maximum with respect to p at a finite (monopoly) price, and that π is strictly increasing in p at all prices below $p_m(c)$ (which denotes the lowest monopoly price at cost c).

Price Competition. In this context, given $r > c$, it is clearly the case that $L(r, c)$ is either the profit accruing to the leader when he limit prices the competitive fringe, or it equals monopoly profits.[9] The latter case arises if $p_m(c)$ is smaller than the limit price r. Solely for ease of exposition, we shall assume throughout that the price-setting equilibrium necessarily involves limit pricing. That is, $r_t \leqslant p_m(c_i)$ for all t.[10] In the other case, the follower's cost is of no consequence to the leader, and nothing of substance is lost by ignoring this alternative.

Quantity Competition. With firms choosing quantities, we assume that the demand curve is such that a unique Cournot–Nash equilibrium exists.[11] $L(r, c)$ then denotes the leader's equilibrium payoff in this game.

We complete our description of the model with a statement of the basic problem faced by the leader. The leader and the follower interact over an infinite time horizon. The leader seeks to maximize the present value of profits, discounted at some rate $\rho > 0$. He must optimally choose the dates $T^1, ..., T^n$ at which to adopt the innovations $1, ..., n$. Each innovation may be adopted instantly, at some future date, or never. Given the ordering of the stages, the leader is constrained to pass through the stages $1, ..., i$ before he can contemplate adopting innovation $i+1$. However, some or all of the innovations may be adopted simultaneously. Of course, the leader is aware of the diffusion process that brings down the costs of the competitive fringe.

[9] It should be pointed out that there are some nontrivial technical issues involved in establishing the validity of this statement. First, in the case that the monopoly price exceeds the limit price, there are problems with the existence of an equilibrium in pure strategies. This problem can be resolved by using a finite grid for prices, or by altering the sharing rule when firms "tie" in prices, or by studying mixed strategies. In all these cases, the payoffs of the limit pricing solution is achieved as an equilibrium. Second, there could be equilibria involving different sets of payoffs. But all these other equilibria involve at least one firm selling (or threatening to sell) below cost. Such strategies may be ruled out on grounds of lack of credibility, and then only the equilibrium payoffs described in the text remain. In this paper, however, we wish to keep these technical issues to a minimum.

[10] A sufficient condition for this is that $r^0 \leqslant p_m(c^n)$.

[11] In Section 5, we impose a further condition on the demand curve to analyze the quantity-setting case.

Formally, given an initial cost vector (r^0, c^0), the leader solves the following dynamic programming problem.

$$\max_{T^1,\dots,T^n} V(r^0, T^1, \dots, T^n) \equiv \int_0^\infty L(r_t, c_t) e^{-\rho t}\, dt - \sum_{i=1}^n e^{-\rho T^i} X^i \tag{11}$$

subject to

$$0 \leqslant T^i \leqslant \infty, \qquad T^{i+1} \geqslant T^i \tag{12}$$

$$\dot{r}_t = -\lambda(r_t - c_t), \qquad r_0 = r^0 \tag{13}$$

$$\begin{aligned} c_t &= c^0, && 0 \leqslant t < T^1 \\ &= c^k, && T^k \leqslant t < T^{k+1} \\ &= c^n, && t \geqslant T^n. \end{aligned} \tag{14}$$

4. Adoption Decisions under Price Competition

We start by defining the value functions for the leader's dynamic programming problem, corresponding to each possible stage of the innovation process. Starting with the last stage n, define for each $r > c^n$,

$$v(r, n) \equiv \int_0^\infty e^{-\rho t} D(r_t)(r_t - c^n)\, dt, \tag{15}$$

where

$$r_t \equiv e^{-\lambda t}(r - c^n) + c^n, \qquad t \geqslant 0. \tag{16}$$

Recursively, having defined $v(r, s)$ for all $s = k+1, \dots, n$, define for each $r > c^k$,

$$v(r, k) \equiv \sup_{T \geqslant 0} \left[\int_0^T e^{-\rho t} D(r_t)(r_t - c^k)\, dt + e^{-\rho T}[v(r_T, k+1) - X^{k+1}] \right], \tag{17}$$

where

$$r_t \equiv e^{-\lambda t}(r - c^k) + c^k, \qquad t \geqslant 0. \tag{18}$$

It is convenient to pose the leader's problem as choosing a "threshold" follower cost at which the innovation will be introduced. The change of variable to z where $(z - c^n)/(r - c^n) = e^{-\lambda t}$ (from (16)) yields for any $r > c^n$

$$v(r, n) = \frac{1}{\lambda} \int_{c^n}^r \left(\frac{z - c^n}{r - c^n} \right)^{\rho/\lambda} D(z)\, dz \tag{19}$$

while $v(r, n) = 0$ if $r = c^n$. A similar change of variable yields the value function at earlier stages

$$v(r, k) = \max_{c^k \leqslant r' \leqslant r} \left\{ \frac{1}{\lambda} \int_{r'}^{r} \left(\frac{z - c^k}{r - c^k} \right)^{\rho/\lambda} D(z)\, dz + \left(\frac{r' - c^k}{r - c^k} \right)^{\rho/\lambda} [v(r', k+1) - X^{k+1}] \right\} \tag{20}$$

if $r > c^k$. If $r = c^k$, then of course

$$v(c^k, k) = \sup_{T \geqslant 0} e^{-\rho T} \{ v(c^k, k+1) - X^{k+1} \}. \tag{21}$$

The following result characterizes the value functions and the associated optimal policy functions.[12]

PROPOSITION 4.1. *The value functions $v(\cdot, k)$, $k = 0, 1, ..., n$ are all well defined, continuous, and strictly increasing in their first argument. In addition, for each $k = 0, 1, ..., n-1$:*

(a) *Given any $r > c^k$, the maximum in (20) is attained at a unique value, which is denoted by $g(r, k)$.*

(b) *There is a threshold value $r^{k+1} \geqslant c^k$ such that the function $g(\cdot, k)$ has the form*

$$g(r, k) = \begin{cases} r & \text{if } c^k \leqslant r \leqslant r^{k+1} \\ r^{k+1} & \text{if } r > r^{k+1}. \end{cases} \tag{22}$$

Thus, conditional on being at the kth stage, the $(k+1)$st innovation is adopted at the first instant that the followers cost is less than or equal to r^{k+1}.

(c) *If $c^k \leqslant r \leqslant r^{k+1}$, then $v(r, k) = v(r, k+1) - X^{k+1}$.*

(d) *$v(\cdot, k)$ is differentiable, with*

$$v_r(r, k) = \begin{cases} v_r(r, k+1) & \text{if } c^k \leqslant r \leqslant r^{k+1} \\ \dfrac{D(r)}{\lambda} - \left(\dfrac{\rho}{\lambda} \right) \dfrac{v(r, k)}{(r - c^k)} & \text{if } r > r^{k+1}. \end{cases} \tag{23}$$

Proof See the Appendix.

Part (b) of this Proposition demonstrates a distinctive feature of the price competition model: the leader is motivated to adopt an innovation

[12] In the case $r = c^k$, the nature of the optimal policy is obvious: conditional on being in stage k adopt the $(k+1)$st innovation instantaneously if $v(c^k, k+1) \geqslant X^{k+1}$ and never otherwise.

only if the follower's cost falls far enough to exert sufficient competitive pressure. In particular if the leader is at stage k, and the follower's current cost is less than or equal to the threshold r^{k+1}, the $(k+1)$th innovation is adopted instantaneously. Otherwise the leader waits for the follower's cost to drop to r^{k+1} before adopting. Of course, if $r^{k+1} = c^k$, then the $(k+1)$th innovation is never adopted.

The sequence of thresholds $r^1, r^2, ..., r^n$ need not be monotone decreasing. This implies that it may be optimal for the leader to group successive innovations; e.g., if $r^k < r^{k+1}$ then innovations k and $(k+1)$ are adopted at the same time, or else neither of them is ever adopted. So innovations will be grouped in the following manner: there are $l \in \{1, 2, ..., n\}$ groups, where group i consists of innovations $F(i), F(i)+1, ..., L(i)$, i.e., $F(i)$ and $L(i)$ denote the first and last innovations in this group. Clearly, $L(i) \geqslant F(i)$ and $L(i)+1 = F(i+1)$; also $F(1)=1$ and $L(l)=n$. The groups are defined by the property that $r^{F(i)} \leqslant r^k$ for $k = F(i)+1, ..., L(i)$ and $r^{F(i)} > r^{F(i+1)} = r^{L(i)+1}$. All of the innovations in group i are adopted when the follower's unit cost falls to the threshold level $r^{F(i)}$. Following this is a period of diffusion while the leader awaits the fall of the follower's cost to the threshold level $r^{F(i+1)}$ of the following group $(i+1)$, at which point this group of innovations is adopted. Of course, if $r^{F(i+1)} = c^{L(i)}$, then the $(i+1)$th group is never adopted. In general, thus, there will be periodic bursts of innovation, followed by intervening periods of diffusion.

One might expect that the precise pattern of grouping of innovations will depend on parameters such as the diffusion rate λ, the discount rate ρ, and the demand function $D(p)$. This, however, turns out to be never true: the groups $\{F(i), L(i); i = 1, ..., l\}$ depend only on the "R & D technology," i.e., the cost levels and adoption costs $\{c^i, X^i;\ i = 1, ..., n\}$ associated with different potential innovations.

PROPOSITION 4.2. *The optimal grouping of innovations* $\{F(i), L(i);$ $i = 1, ..., l\}$ *depends only on* $\{c^i, X^i;\ i = 1, ..., n\}$, *the unit cost levels, and adoption costs associated with successive innovations.*

Proof. See the Appendix.

In what follows, we shall examine the effects of varying the diffusion rate λ on the leader's optimal innovation strategy. Since Proposition 4.2 implies that the grouping of innovations is independent of λ, we may proceed on the assumption that we are dealing with a set of l (compound) innovations, where each compound innovation corresponds to an optimal grouping of primary innovations. To preserve notational uniformity, we shall henceforth assume that there are l innovations, with c^i, X^i, and r^i being the unit production cost, adoption cost and threshold follower cost associated with the ith innovation. Since the innovations are already optimally grouped, it follows that the thresholds r^i form a monotonic sequence, $r^1 > r^2 > \cdots > r^I$, and an innovation is adopted singly if the follower cost r becomes equal to the corresponding threshold.

We begin by describing the total number of innovations that will optimally be adopted (in finite time) by the leader. Proposition 4.1 implies that conditional on the $(i-1)$th innovation being adopted, the ith innovation will be adopted if and only if $r^i > c^{i-1}$. This, in turn, is equivalent to the condition that $v(c^{i-1}, i) - X^i > 0$.[13] So the total number of innovations adopted is given by N where $v(c^{j-1}, j) - X^j > 0$ for $j = 1, ..., N$ and $v(c^N, N+1) - X^{N+1} \leqslant 0$.

Now consider an increase in the diffusion rate λ. Given any stage j, this lowers the cost r_t of the follower at all $t > 0$. Given our assumption that $r_t < p_m(c_t)$, and that $D(p)(p-c)$ is strictly increasing in p for all prices p below the corresponding monopoly price, it follows (from a straightforward induction argument) that an increase in λ reduces $v(r, j)$ for all $r > c^{j-1}$. This establishes the following:

PROPOSITION 4.3. *An increase in the diffusion rate λ cannot increase the number of innovations adopted.*[14]

Thus, the effects of increased diffusion on the total number of innovations adopted are quite Schumpetarian: fewer innovations are adopted because faster diffusion erodes the profitability of adopted innovations. However, this may provide a misleading picture of the overall effect of diffusion on innovation incentives. In a dynamic setting innovative activity is necessarily multidimensional; in particular, the effect on the *timing* of adoption decisions is also important.

In fact, the picture is considerably different once we examine the effects of diffusion on the timing of innovations. Without loss of generality, assume that all innovations are adopted for λ small enough.[15] Now note that the threshold $r^k(\lambda)$ for the kth innovation at diffusion rate λ, is decreasing in λ.[16] That is, if the diffusion rate increases, then a lower level of follower cost is required to exert enough "competitive pressure" on the

[13] This follows from the fact that r^i solves $Q(r^i, i-1) = ((c^{i-1} - c^i)/(r^i - c^i))\, v(r^i, i) - X^i = 0$, and that Q is strictly decreasing (see the Appendix, Proof of Proposition 4.1); so $r^i > c^{i-1}$ is equivalent to $Q(c^{i-1}, i-1) = v(c^{i-1}, i) - X^i > 0$.

[14] This presumes that the diffusion rate is strictly positive to begin with. The result may not hold if λ goes from zero to positive (see Mookherjee and Ray [12]).

[15] If some innovation is not adopted for λ small enough, then it (and all succeeding potential innovations) may be ignored, since the preceding arguments imply that such an innovation will not be adopted for *any* diffusion rate.

[16] This follows from the fact that $r^k(\lambda)$ solves $Q(r, k-1, \lambda) = ((c^{k-1} - c^k)/(r - c^k))\, v(r, k, \lambda) - X^k = 0$, where λ is used as an explicit argument in v and Q. Since an increase in λ lowers v, and Q is decreasing in r, $r^k(\lambda)$ must be decreasing.

leader to innovate at any stage. Intuitively, this follows from the fact that the leader's innovation incentive is a decreasing function both of the diffusion rate, and the follower's cost level. Now define $\tilde{r}^1 = \lim_{\lambda \downarrow 0} r^1(\lambda)$, and assume that $r^0 > \tilde{r}^1$. In words, for small values of λ, we suppose that the leader will not adopt the first innovation instantaneously. In what follows, we consider the timing $T(\lambda)$ of the first innovation, and how it varies with λ.

PROPOSITION 4.4. *Suppose that $r^0 > \tilde{r}^1$; i.e., for small enough values of the diffusion rate λ, the first innovation is not adopted at $t = 0$. Then there exists $\lambda^* > 0$ such that for any $\lambda \in (0, \lambda^*)$, the timing $T(\lambda)$ of the first innovation is finite, but $T(\lambda)$ goes to infinity as λ tends either to 0 or λ^* in this interval.*

The proof of this is as follows. Let λ^* be defined by $v(c^0, 0, \lambda^*) - X^1 = 0$, so the first innovation is adopted in finite time if and only if $\lambda < \lambda^*$.[17] For any $\lambda \in (0, \lambda^*)$, the adoption time $T(\lambda)$ satisfies $r^1(\lambda) = c^0 + e^{-\lambda T(\lambda)}[r^0 - c^0]$. Hence, $r^0 > \tilde{r}^1$ implies that $T(\lambda) \to \infty$ as $\lambda \to 0+$. Also, $r^1(\lambda^*) = c^0$, so $T(\lambda) \to \infty$ as $\lambda \to \lambda^*-$.

Proposition 4.4 serves to establish the basic point of this paper: the relation between diffusion and innovation dates is not necessarily monotonic, and both "Leibensteinian" and "Schumpetarian" effects may therefore coexist. The intuition is as follows. By Proposition 4.1, the timing of the first innovation is determined by the date at which the follower's cost drops to the threshold value $r^1(\lambda)$. This date depends on the speed of diffusion in two ways. First, the threshold value $r^1(\lambda)$ is decreasing in λ: an increased diffusion rate then requires the follower's cost to fall to a lower level before triggering the first innovation. This delays the innovation date, and represents the Schumpetarian effect. Second, the speed at which the follower's cost drops towards the threshold value is determined directly by the diffusion rate. Faster diffusion causes the follower's cost to drop faster towards the leader's cost, thereby causing the leader to innovate earlier. This is the Leibensteinian effect.

When there is a very "small amount" of diffusion, the first innovation is indeed adopted, but only in the distant future.[18] The reason is that it takes a long time for the follower's cost to fall sufficiently to exert the necessary amount of competitive pressure on the leader: the Leibensteinian effect dominates over this range. On the other hand, when λ approaches λ^* (from below), the threshold value $r^1(\lambda)$ approaches the leader's cost c^0. This is

[17] λ^* exists since $v(c^0, 0, \lambda) \to 0$ as $\lambda \to \infty$, while by assumption the first innovation is adopted for λ small enough. The continuity of v in λ follows from an induction argument.

[18] A similar result holds for the adoption of later innovations as well. An analogous argument establishes that as $\lambda \to 0$, an adjacent pair of (groups of) innovations will have an increasingly larger time period separating their adoption.

because the rents from the innovation are now tending to vanish. Despite the fact that the follower's cost is now falling more rapidly, it takes, again, a long time for it to fall to such a low threshold. Over this range, then, the Schumpetarian effect dominates.

Proposition 4.4 leaves the intermediate behavior of the adoption time unspecified. In general, we have not been able to characterize this behavior, though there is reason to conjecture that the function will be U-shaped. This conjecture is borne out in the example studied in Mookherjee and Ray [12], where a linear demand curve permits explicit computation.

5. Adoption Decisions Under Quantity Competition

It turns out that the results of the previous section are substantially altered once the product market is characterized by quantity competition. Our main result (Proposition 5.1) states that all available innovations that are *ever* adopted will be adopted instantly.

We assume that for each pair (r, c), quantity competition yields a unique equilibrium payoff vector, and that the quantities produced by the leader and the follower (Q_1 and Q_2, respectively) are strictly positive.[19]

Our proof makes use of the following conditions imposed on the quantity-setting game at any date t. Let $P(\cdot)$ denote the inverse demand curve.

(C1) $P(Q)$ is thrice differentiable whenever $P > 0$, and $P''(Q) \leqslant 0$.

By (C1), the derivatives $L_r(r, c)$, $L_{rr}(r, c)$ are well defined. We also assume

(C2) $L_{rr}(r, c) \leqslant 0$ for all $r \geqslant c$.

The economic content of (C1) is essentially that the demand curve is concave. We do not defend it except to say that it is an assumption frequently made in the literature. Condition (C2) *is* a stronger restriction. It says that the effect of an increase in follower cost on the leader's payoff is diminishing as the follower's cost rises. While we do not find this an unreasonable assumption, it is nevertheless nontrivial.[20]

It can be checked that a linear demand curve, or more generally, any

[19] The uniqueness of Cournot outputs will be a consequence of condition (C1) below on the demand curve.

[20] Note that other analyses of R & D often have to invoke special assumptions on the nature of Cournot–Nash profit functions, e.g., Reinganum [15]. However, her assumptions are somewhat weaker, essentially requiring that $L_{rc} < 0$, a condition implied by (C1) and (C2). The condition $L_{rc} < 0$ plays an important role in the proof of Proposition 5.1; the intuitive explanation of this has been provided in Section 2.

demand curve of the form $P(Q) = A - bQ^c$, for $A > 0$, $b > 0$, $c \geq 1$, satisfies both (C1) and (C2).

We are now in a position to state

PROPOSITION 5.1. *Suppose that the demand curve satisfies conditions* (C1) *and* (C2). *Then, under quantity competition, all the innovations that are adopted at all are adopted instantly. Moreover, if a certain set of innovations is adopted at some initial follower cost r, then this set will continue to be adopted for all* $r' > r$.

So, under conditions (C1) and (C2), quantity competition exhibits markedly different features from its price-setting counterpart. An existing stock of innovations is never "spaced apart" in the adoption process; adoption occurs instantly, or never.

Put another way, Proposition 5.1 states that the diffusion of technical knowledge to the follower creates no additional competitive pressure. Indeed, the proof of the proposition reveals that under quantity competition, the leader has a *greater* incentive to innovate when the follower has a *higher* cost (i.e., L_{rc} is negative). Therefore, the existing stock of technical knowledge that is potentially profitable is actually profitable at the initial date itself. In our quantity competition model, cyclical waves of innovation and diffusion will coincide exactly with the spurts in the development of applicable scientific knowledge. In contrast, price competition creates (via the diffusion process) conditions for further innovation that did not exist when follower costs were high.

Proposition 5.1 also states that innovation is more likely when the initial disparities between leader and follower are *high* rather than low, which is in marked contrast to the price competition model.

Observe that under quantity competition, the innovation process is not Schumpetarian, in the sense that there are no cycles of alternating diffusion and innovation. On the other hand, it is Schumpetarian in that an increase in the rate of diffusion can only lower the commercial viability of innovations:

PROPOSITION 5.2. *Assume* (C1) *and* (C2). *Under quantity competition, an increase in the rate of diffusion cannot increase the number of innovations to be adopted, and will generally lower it.*

APPENDIX

Proof of Proposition 4.1. Equation (19) implies that $v(\cdot, n)$ is well-defined, continuous, and strictly increasing over $r \geq c^n$. Arguing by induction, it follows from (20) that each of the functions $v(\cdot, k)$, $k = 0, 1, ..., n-1$

is well-defined (since a continuous function is being maximized over a compact set) and continuous (by the Maximum Theorem) at any $r > c^k$. Continuity at $r = c^k$ also follows from the fact that this is true for $v(\cdot, n)$, and by an inductive argument. That $v(\cdot, k)$ is strictly increasing follows from the argument that given any r and r' with $r > r' \geq c^k$, the leader always has the option (when the follower's cost is r) of not innovating as long as $r > r'$, and choosing the optimal policy from r' onwards when the follower's cost becomes equal to r'. Since the leader will earn positive profits in the first phase, it must be the case that $v(r, c^k) > v(r', c^k)$.

That (a)–(d) hold is established by induction. We first show that it holds for $k = (n-1)$. Let $S(r')$ denote $[(1/\lambda) \int_{r'}^{r} ((z - c^{n-1})/(r - c^{n-1}))^{\rho/\lambda} D(z)\, dz + ((r' - c^{n-1})/(r - c^{n-1}))^{\rho/\lambda} \{v(r', n) - X^n\}]$ for any r' between r and c^{n-1}, so $v(r, n-1) \equiv \max_{c^{n-1} \leq r' \leq r} S(r')$. Now (19) implies that $v(r, n)$ is differentiable, with $v_r(r, n) = D(r)/\lambda - \rho v(r, n)/\lambda(r - c^n)$. Hence $S(r')$ is differentiable; letting $\psi(r')$ denote the derivative of S, we obtain

$$\psi(r') = \frac{\rho}{\lambda} \frac{(r' - c^{n-1})^{(\rho/\lambda) - 1}}{(r - c^{n-1})^{\rho/\lambda}} \cdot Q(r', n), \tag{24}$$

where $Q(r, n) \equiv ((c^{n-1} - c^n)/(r - c^n))\, v(r, n) - X^n$, for any $r > c^n$. Clearly, $Q(\cdot, n)$ is differentiable. We claim $Q(\cdot, n)$ is strictly decreasing. To establish this, note that sign $Q_r(r, n) = \text{sign}[v_r(r, n) - (v(r, n)/(r - c^n))]$ and

$$\begin{aligned} \frac{v(r, n)}{r - c^n} - v_r(r, n) &= \frac{(\lambda + \rho)\, v(r, n)}{\lambda(r - c^n)} - \frac{D(r)}{\lambda} \\ &= \frac{(\lambda + \rho)}{\lambda^2 (r - c^n)} \int_{c^n}^{r} \left(\frac{z - c^n}{r - c^n}\right)^{\rho/\lambda} D(z)\, dz - \frac{D(r)}{\lambda} \\ &> \frac{D(r)}{\lambda} \left[\frac{\lambda + \rho}{\lambda(r - c^n)} \int_{c^n}^{r} \left(\frac{z - c^n}{r - c^n}\right)^{\rho/\lambda} dz - 1\right] = 0, \end{aligned} \tag{25}$$

where the inequality uses the fact that the demand curve is downward sloping.

By (24), the sign of $\psi(r')$ is the same as that of $Q(r', n)$. Since $\lim_{r' \to \infty} Q(r', n) < 0$, (a) and (b) follow from the argument that *either* (i) $\lim_{r' \downarrow c^{n-1}} Q(r', n) \leq 0$, in which case $Q(r, n) < 0$ for all $r > c^{n-1}$, so $g(r, n-1) = c^{n-1}$ and it pays to never adopt the nth innovation, *or* (ii) $\lim_{r' \downarrow c^{n-1}} Q(r', n) > 0$, in which case there exists a unique r^n such that $Q(r^n, n) = 0$. In this case the fact that $Q(\cdot, n)$ is strictly decreasing implies that $\min(r, r^n)$ is the unique maximizer of S with respect to r' between r and c^{n-1}.

For $k = n-1$, (c) is obvious, and it remains to show that (d) holds.

Parts (b) and (c) imply that $v(\cdot, n-1)$ is differentiable at $r \neq r^n$, and differentiation yields

$$v_r(r, n-1) = \begin{cases} \dfrac{D(r)}{\lambda} - \dfrac{\rho v(r, n-1)}{\lambda(r - c^{n-1})} & \text{if } r > r^n \\ \dfrac{D(r)}{\lambda} - \dfrac{\rho v(r, n-1)}{\lambda(r - c^{n-1})} + \dfrac{\rho Q(r, n)}{\lambda(r - c^{n-1})} & \text{if } r < r^n. \end{cases}$$

Since $Q(r^n, n) = 0$, it follows that $\lim_{r \downarrow r^n} v_r(r, n-1) = \lim_{r \uparrow r^n} v_r(r, n-1)$. So v is differentiable at $r = r^n$, as well, and (d) is established.

Next, we show that if (a)–(d) hold for $k = m, ..., n$ they also hold for $k = m-1$. Let $\psi(r', r, m-1)$ denote the derivative of the right hand side of (20), defined for r' between c^{m-1} and r. Then,

$$\psi(r', r, m-1) = \left(\frac{r' - c^{m-1}}{r - c^{m-1}}\right)^{\rho/\lambda} \left[-\frac{D(r')}{\lambda} + v_r(r', m) + \frac{\rho}{\lambda} \frac{v(r', m) - X^m}{r' - c^{m-1}}\right]. \tag{26}$$

Suppose $r' \leqslant r^{m+1}$. Choose $M \in \{m+1, ..., n\}$ such that $r' \leqslant r^k$ for $k = m+1, ..., M$ and $r' > r^{M+1}$. Since (c) and (d) hold for $k = m, ..., n$, $v(r', m) = v(r', M) - \sum_{i=m+1}^{M} X^i$, and $v_r(r', m) = v_r(r', M) = D(r')/\lambda - (\rho/\lambda)(v(r', M)/(r' - c^M))$ by virtue of (23). Using this in (25), we obtain (combining with a straightforward computation for the case $r > r^{m+1}$ which uses expression (d) for $v_r(r', m)$)

$$\psi(r', r, m-1) = \frac{\rho}{\lambda}\left(\frac{r' - c^{m-1}}{r - c^{m-1}}\right)^{\rho/\lambda} \frac{Q(r', m-1)}{r' - c^{m-1}},$$

where

$$Q(r', m-1) = \begin{cases} \dfrac{(c^{m-1} - c^m)\, v(r', m)}{(r' - c^m)} - X^m & \text{if } r' > r^{m+1} \\ \dfrac{(c^{m-1} - c^M)\, v(r', M)}{r' - c^M} - \displaystyle\sum_{i=m}^{M} X^i & \text{otherwise.} \end{cases} \tag{27}$$

Properties (a)–(c) now follow by reasoning analogous to the case of $k = n-1$, i.e., that $Q(\cdot, m-1)$ is continuous and strictly decreasing at any $r > c^{m-1}$. The continuity of Q is obvious at $r' \neq r^{m+1}$. To establish continuity at $r' = r^{m+1}$, note that the right hand limit is $((c^{m-1} - c^m)/(r^{m+1} - c^m))\, v(r^{m+1}, m) - X^m$, while the left hand limit is $((c^{m-1} - c^M)/(r^{m+1} - c^M))\, v(r^{m+1}, M) - \sum_{i=m}^{M} X^i$. Now by (c), $v(r', m) = v(r', M) - \sum_{i=m+1}^{M} X^i$ for any $r' \leqslant r^{m+1}$. This implies that $v(r', M)/(r' - c^M) = v(r', m)/(r' - c^m) - Q(r', m)/(r' - c^m)$ for any $r' \leqslant r^{m+1}$.

Since r^{m+1} is defined by the property $Q(r^{m+1}, m) = 0$, it follows that $v(r^{m+1}, M)/(r^{m+1} - c^M) = v(r^{m+1}, m)/(r^{m+1} - c^m)$. Hence the left hand limit exceeds the right hand limit by $Q(r^{m+1}, m)$, which equals 0, and Q is continuous at $r' = r^{m+1}$.

To establish that $Q(\cdot, m-1)$ is strictly decreasing, the continuity of the function implies that it suffices to check that it is strictly decreasing at all $r' \neq r^s$, $s = m, m+1, ..., n$; $r' \geqslant c^{m-1}$. At all such values of r', Q is differentiable. Using the convention that $M = m$ if $r' > r^{m+1}$, the derivative of Q is

$$Q_{r'} = \left(\frac{c^{m-1} - c^M}{r' - c^M}\right)\left[v_r(r', M) - \frac{v(r, M)}{r' - c^M}\right] \tag{28}$$

at $r' \neq r^s$, $s = m, m+1, ..., n$, since at such points, M is locally independent of r'. Now by (23) and the induction hypothesis, $v_r(r', M) = D(r')/\lambda - (\rho/\lambda)(v(r', M)/(r' - c^M))$, so that $v_r(r', M) - (v(r', M)/(r' - c^M)) = D(r')/\lambda - ((\rho + \lambda)/\lambda)(v(r', M)/(r' - c^M))$. Since at the Mth stage the leader always has the option of not adopting the $(M+1)$th innovation, it follows that the right hand side of (28) is not larger than

$$\frac{D(r')}{\lambda} - \frac{(\rho + \lambda)}{\lambda^2(r' - c^M)} \int_{c^M}^{r'} \left(\frac{z - c^M}{r' - c^M}\right)^{\rho/\lambda} D(z)\, dz$$

which is negative via reasoning analogous to (25), since $r' \geqslant c^{M-1} > c^M$. This completes the proof that (a)–(c) hold for $k = m-1$; the proof of (d) is identical to that used for $k = n-1$. ∎

Proof of Proposition 4.2. Given any group i, it is true that $r^{F(i)} \leqslant r^k$ for all $k = F(i)+1, ..., L(i)$, and $r^{F(i)} < r^k$ for $k \in \{1, ..., F(i)-1\}$. (The latter inequality follows from the fact that $r^{F(i)} < r^{F(i-S)} \leqslant r^k$ for any group $(i-S) < i$ and any innovation k in this group.)

Take any $k \in \{1, ..., F(i)-1\}$. Since $Q(\cdot, k-1)$ is strictly decreasing, $Q(r^{F(i)}, k-1) > Q(r^k, k-1) = 0 = Q(r^{F(i)}, F(i)-1)$. Now $Q(r^{F(i)}, k-1) = ((c^{k-1} - c^{L(i)})/(r^{F(i)} - c^{L(i)}))\, v(r^{F(i)}, L(i)) - \sum_{j=F(i)}^{L(i)} X^j$, from which it follows that

$$(c^{k-1} - c^{L(i)})/\sum_{j=k}^{L(i)} X^j > (c^{F(i)-1} - c^{L(i)})/\sum_{j=F(i)}^{L(i)} X^j. \tag{29}$$

An analogous argument establishes the weak inequality counterpart of (29) for $k \in \{F(i)+1, ..., L(i)\}$. Hence, given the last innovation $L(i)$ in a group, and defining the function $\alpha(k, L(i)) \equiv (c^{k-1} - c^{L(i)})/\sum_{j=k}^{L(i)} X^j$ for $k \leqslant L(i)$, the first innovation $F(i)$ of the group equals $\min\{\arg\min_{k \leqslant L(i)} \alpha(k, L(i))\}$. Given $L(l) = n$ and working backwards recursively, this uniquely identifies the l groups. ∎

For the proofs of Propositions 5.1 and 5.2, recall that we are using $L(r, c)$ to denote the Cournot payoffs to the leader when the cost configuration is (r, c). Define, for $r > c^n$

$$v(r, n) \equiv \int_0^\infty e^{-\rho t} L(r_t^n, c^n)\, dt, \tag{30}$$

where

$$r_t^n \equiv c^n + e^{-\rho t}(r - c^n), \qquad t \geqslant 0. \tag{31}$$

And recursively, having defined $v(r, s)$ for $r > c^s$ and $s = k+1, ..., n$, define for stage k and $r > c^k$,

$$v(r, k) \equiv \sup_{T \geqslant 0} \left\{ \int_0^T e^{-\rho t} L(r_t^k, c^k)\, dt + e^{-\rho T}[v(r_T^k, k+1) - X^{k+1}] \right\}, \tag{32}$$

where

$$r_t^k \equiv c^k + e^{-\lambda t}(r - c^k), \qquad t \geqslant 0. \tag{33}$$

Proof of Proposition 5.1. Suppose, first, that $n = 1$. Define for $r > c^0$,

$$E(r) \equiv \int_0^\infty e^{-\rho t}[L(r_t^1, c^1) - L(r_t^0, c^0)]\, dt - X^1. \tag{34}$$

Using (30)–(33) for $k = 0$ and $n = 1$, we have for $r > c^0$,

$$v(r, 0) \equiv \int_0^\infty e^{-\rho t} L(r_t^0, c^0)\, dt + \sup_{T \geqslant 0} e^{-\rho T} E(r_T). \tag{35}$$

The result for $n = 1$ follows if $E(r)$ is a nondecreasing function of r. To show this, it suffices (from (34)) to prove that for each r and each $t \geqslant 0$, $Z_t(r) \equiv L(r_t^1, c^1) - L(r_t^0, c^0)$ is a nondecreasing function of r. Now,

$$\frac{dZ_t(r)}{dr} = e^{-\lambda t}[L_r(r_t^1, c^1) - L_r(r_t^0, c^0)]. \tag{36}$$

Now note that $r_t^0 \geqslant r_t^1 \geqslant c^1$ for all t, so using condition (C2), it will suffice to show that for all t,

$$L_r(r_t^0, c^1) \geqslant L_r(r_t^0, c^0), \tag{37}$$

which in turn will follow if $L_{rc}(r, c) \leqslant 0$ whenever $r - c \geqslant 0$.

For any such pair (r, c), use Q_1 and Q_2 to denote the equilibrium leader and follower outputs, respectively, and let $q = Q_1 + Q_2$. By the Envelope Theorem, $L_r(r, c) = Q_1 P'(q)(\partial Q_2/\partial r)$, so that

$$L_{rr}(r, c) = P'(q) \frac{\partial Q_2}{\partial r} \frac{\partial Q_1}{\partial r} + Q_1 \frac{\partial}{\partial r}\left[P'(q) \frac{\partial Q_2}{\partial r} \right] \tag{38}$$

and

$$L_{rc}(r, c) = P'(q) \frac{\partial Q_2}{\partial r} \frac{\partial Q_1}{\partial c} + Q_1 \frac{\partial}{\partial c}\left[P'(q) \frac{\partial Q_2}{\partial r} \right]. \tag{39}$$

Routine computation shows that

$$P'(q) \frac{\partial Q_2}{\partial r} = \frac{2 + \eta(q) x}{3 + \eta(q)},$$

where $\eta(q) \equiv qP''(q)/P'(q)$ and $x \equiv Q_1/q$. Consequently,

$$\frac{\partial}{\partial r}\left[P'(q) \frac{\partial Q_2}{\partial r} \right] = \frac{\eta(q)}{3 + \eta(q)} \frac{\partial x}{\partial r} + \frac{\partial q}{\partial r}\left[\frac{\partial}{\partial q}\left\{ \frac{2 + \eta(q) x}{3 + \eta(q)} \right\} \right] \tag{40}$$

while

$$\frac{\partial}{\partial c}\left[P'(q) \frac{\partial Q_2}{\partial r} \right] = \frac{\eta(q)}{3 + \eta(q)} \frac{\partial x}{\partial c} + \frac{\partial q}{\partial c}\left[\frac{\partial}{\partial q}\left\{ \frac{\eta(q) x}{3 + \eta(q)} \right\} \right]. \tag{41}$$

Now (40) exceeds (41), since the first order conditions for a Cournot equilibrium imply that $\partial q/\partial c = \partial q/\partial r$, $\partial x/\partial r > 0 > \partial x/\partial c$, while $\eta(q) \geqslant 0$ by (C1). Therefore, $0 \geqslant L_{rr} \geqslant L_{rc}$, as $P'(q)(\partial Q_2/\partial r)(\partial Q_1/\partial r) \geqslant 0 \geqslant P'(q)(\partial Q_2/\partial r)(\partial Q_1/\partial c)$, and the proposition is proved for $n = 1$.

Now use induction to establish the proposition for an arbitrary number of innovations. For any stage k, suppose the proposition is true for stages $k+1, ..., n-1$. Then

$$v(r, k) \equiv \int_0^\infty e^{-\rho t} L(r_t, c^k)\, dt + \sup_{T \geqslant 0} e^{-\rho T} E(r_T, k), \tag{42}$$

where

$$E(r, k) \equiv \int_0^\infty e^{-\rho t}[L(r'_t, c^{s(r)}) - L(r_t, c^k)]\, dt - \sum_{j=k+1}^{s(r)} X^j, \tag{43}$$

$$r'_t \equiv c^{s(r)} + e^{-\lambda t}(r - c^{s(r)}), \qquad t \geqslant 0, \tag{44}$$

$$r_t \equiv c^k + e^{-\lambda t}(r - c^k), \qquad t \geqslant 0, \tag{45}$$

and $s(r)$ is the maximal index (not less than $k+1$) such that innovations $k+1, ..., s(r)$ are adopted instantly, conditional on $k+1$ being adopted at follower cost r. Since the proposition is true from stage $k+1$ onwards, $s(r)$ is nondecreasing. Hence there are at most a finite number of values of r at which $s(r)$ switches value. Now, the Maximum Theorem implies that $E(r, k)$ is continuous in r, since

$$E(r, k) = v(r, k+1) - \int_0^\infty e^{-\rho t} L(r_t, c^k)\, dt - X^{k+1}. \tag{46}$$

Now at any r where $s(r)$ is locally independent of r, an argument identical to that used in the case $n=1$ establishes that $E(r, k)$ is locally nondecreasing. Combining with continuity, it follows that $E(r, k)$ is everywhere nondecreasing on $r > c^k$. Hence the optimal date of innovation k is $T=0$ (if $E(r, k) \geqslant 0$) or $T = \infty$ otherwise. This completes the proof. ▌

Proof of Proposition 5.2. Include λ explicitly in expression (42) to write

$$E(r, k; \lambda) \equiv \int_0^\infty e^{-\rho t}[L(r'_t, c^{s(r,\lambda)}) - L(r_t, c^k)]\, dt - \sum_{j=k+1}^{s(r,\lambda)} X^j, \tag{47}$$

where r'_t and r_t are given by analogues of (44) and (45).

First suppose that $k = n-1$. Then $s(r, \lambda) = n$ for all (r, λ). Differentiating (47) w.r.t. λ, we have

$$E_\lambda(r, n; \lambda) = -\int_0^\infty te^{-(\rho+\lambda)t}[L_r(r'_t, c^n)(r - c^n) - L_r(r_t, c^{n-1})(r - c^{n-1})]\, dt.$$

Recalling that the right hand side of (36) is nonnegative, it follows that

$$E_\lambda(r, n; \lambda) \leqslant 0. \tag{48}$$

Recalling that innovation n takes place if and only if $E(r, n; \lambda) \geqslant 0$, the proposition is established for the last stage. Now use induction to establish the result for all k: an argument similar to the case where $k = n-1$ establishes that $E_\lambda(r, k; \lambda) \leqslant 0$ at any λ where $s(r, \lambda)$ is locally independent of λ, and for other values of λ we use the continuity of E in λ. ▌

References

1. M. Abramovitz, Resource and output trends in the United States since 1870, *Amer. Econ. Rev.* **46** (1956), 5–23.
2. K. J. Arrow, Economic welfare and the allocation of resources for invention, *in* "The Rate and Direction of Inventive Activity: Economic and Social Factors" (R. Nelson, Ed.). Princeton Univ. Press, Princeton, 1962.

3. P. Dasgupta and J. Stiglitz, Industrial structure and the nature of innovative activity, *Econ. J.* **90** (1980), 266–293.
4. P. Dasgupta and J. Stiglitz, Uncertainty, industrial structure and the speed of R & D, *Bell J. Econ.* **11** (1980), 1–28.
5. E. Denison, "Trends in American Economic Growth, 1929–1982," The Brookings Institution, Washington, DC, 1985.
6. C. Futia, Schumpetarian competition, *Quart. J. Econ.* **94** (1980), 675–696.
7. M. Kamien and N. Schwartz, "Market Structure and Innovation," Cambridge Univ. Press, Cambridge, 1982.
8. M. Kamien and N. Schwartz, Timing of innovations under rivalry, *Econometrica* **40** (1972), 43–60.
9. M. Kamien and N. Schwartz, On the degree of rivalry for maximum innovative activity, *Quart. J. Econ.* **90** (1976), 245–260.
10. H. Leibenstein, Allocative efficiency versus X-efficiency, *Amer. Econ. Rev.* **56** (1966), 392–415.
11. G. Loury, Market structure and innovation, *Quart. J. Econ.* **93** (1979), 395–410.
12. D. Mookherjee and D. Ray, "On the Competitive Pressure Induced by the Diffusion of Innovations," Research Paper No. 1009, Graduate School of Business, Stanford University, 1988.
13. R. Nelson and S. Winter, "An Evolutionary Theory of Economic Change," Harvard Univ. Press, Cambridge, MA, 1982.
14. J. Reinganum, Innovation and industry evolution, *Quart. J. Econ.* **100** (1985), 81–100.
15. J. Reinganum, Technology adoption under imperfect information, *Bell J. Econ.* **12** (1983), 618–624.
16. J. Schumpeter, "The Theory of Economic Development," Harvard Univ. Press, Cambridge, MA, 1961.
17. J. Schumpeter, "Capitalism, Socialism and Democracy," Harper, New York, 1950.
18. R. Solow, Technical change and the aggregate production function, *Rev. Econ. Statist.* **39** (1957), 312–320.
19. M. Spence, Cost reduction, competition and industry performance, *Econometrica* **52** (1984), 101–122.